13/06/2025 TE CSE-AIML SEM-V C-SCHEME SAIDS QP CODE: 10083339

Duration: (3 hrs.) **[Maximum Marks : 80]**

NB:

(1) Question No. 1 is compulsory.
(2) Attempt any three questions out of the remaining five.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

Q1. ATTEMPT ANY FOUR [20]

a. What is hypothesis testing ? Explain type I and type II errors?

b. What is Fisher's exact test?

c. Explain the difference between Stratified and Cluster Sampling.

d. Explain Linear Regression and its Applications.

e. Define standard deviation and interquartile range with examples.

Q2. a. Find the correlation coefficient from the given data. [10]

Subject	Experience (X)	Salary (Y)
1	5	50
2	8	60
3	12	75
4	15	85
5	18	95
6	20	105

b. What is Chi-Square Test? A retail company wants to determine if there is a significant association between customer gender and preference for online shopping vs. in-store shopping. The company collected data from a random sample of 200 customers, and the results are summarized in the following contingency table. Use the Chi-Square Test for Independence to determine if there is a statistically significant association between gender and shopping preference at a 5% significance level (α=0.05) [10]

Gender	Prefers Online Shopping	Prefers In-Store Shopping	Total
Male	60	40	100
Female	70	30	100
Total	130	70	200

Q3. a. Explain the concept of p-value in hypothesis testing [10]

b A school conducted an aptitude test for three different grades (Grade A, Grade B, and Grade C). The scores obtained by the students in each grade are given. At a 95% confidence level, determine if the scores differ significantly across the three grades using the Kruskal-Wallis test. [10]

Grade A	Grade B	Grade C
85	78	90
88	82	85
80	75	88
92	85	92
78	80	84
85	88	80

Q4. a. A researcher is analyzing the test scores of students. The sample mean score for 20 students is 250, the expected (population) mean is 260, and the standard deviation is 40. Calculate the z-score for this sample mean. [10]

b. A researcher conducted a survey of 50 college students to determine how many hours they spend studying per week. Create a frequency distribution table for the data provided. [10]

12, 15, 8, 10, 20, 7, 13, 18, 9, 11, 14, 16, 6, 12, 15, 19, 10, 8, 13, 17, 11, 14, 9, 12, 16, 7, 15, 18, 10, 13, 16, 9, 11, 14, 8, 12, 17, 10, 15, 19, 6, 13, 18, 11, 14, 9, 12, 16, 10, 15

Q5. a. A pharmaceutical company has developed a new drug that they claim lowers blood pressure more effectively than the current standard drug. The average reduction in blood pressure for patients using the standard drug is 10 mmHg, with a standard deviation of 5 mmHg. The company conducts a clinical trial with 30 patients using the new drug and observes an average reduction of 12 mmHg. At a 0.05 significance level, answer the following: [10]

1. State the null and alternative hypotheses.
2. Calculate the test statistic.

Determine if the new drug is statistically significantly more effective than the standard drug.

b. Find the simple linear regression equation for the given data. [10]

Time	Growth
3	12
6	18
9	25
12	32
15	40
18	45

Q6. a. Explain the concept of two-way ANOVA. How does it differ from one-way ANOVA? Describe the assumptions of two-way ANOVA and how you would check these assumptions. Also, briefly explain Friedman's test as a non-parametric alternative. [10]

b. Write short notes on (any two) [10]

1. Chi-square distribution.
2. Weibull distribution.
3. Stem & Leaf Plot
4. Box Plot.

****